Anonymous Communication and Cell Counting Attack Against Tor

K. Rajesh Kumar, B. Praveen

Dept. of CSE, Swarnandhra College of Engg & Tech, Narsapuram, AP, India

Abstract
Various low-latency anonymous communication systems such as Tor and Anonymousizer have been designed to provide anonymity service for users. In order to hide the communication of users, most of the anonymity systems pack the application data into equal-sized cells. Via extensive experiments on Tor, we found that the size of IP packets in the Tor network can be very dynamic because a cell is an application concept and the IP layer may repack cells. Based on this finding, we investigate Anonymous Communication and Cell Counting Attack against Tor, which allows the attacker to confirm anonymous communication relationship among users very quickly. In this attack, by marginally varying the number of cells in the target traffic at the malicious exit onion router, the attacker can embed a secret signal into the variation of cell counter of the target traffic. The embedded signal will be carried along with the target traffic and arrive at the malicious entry onion router. Then, an accomplice of the attacker at the xmalicious entry onion router will detect the embedded signal based on the received cells and confirm the communication relationship among users. We have implemented this attack against Tor, and our experimental data validate its feasibility and effectiveness. There are several unique features of this attack. First, this attack is highly efficient and can confirm very short communication sessions with only tens of cells. Second, this attack is effective, and its detection rate approaches 100% with a very low false positive rate. Third, it is possible to implement the attack in a way that appears to be very difficult for honest participants to detect.

Keywords
Anonymous Communication, Secret Spread-Spectrum Signal, Traffic Flow, Tor Network

I. Introduction
Now a day’s internet usage is rapidly increasing based on the increasing of the internet the privacy and security have received greater attention with the rapid growth and public acceptance of the Internet, which has been used to create our global E-economy. Anonymity has become a necessary and legitimate aim in many applications, including anonymous Web browsing, Location-Based Services (LBSs), and E-voting. In these applications, encryption alone cannot maintain the anonymity required by participants [1]–[3]. Research on flow-based anonymity applications has recently received great attention in order to preserve anonymity in low-latency applications, including Web browsing and peer-to-peer file sharing [5–6]. To degrade the anonymity service provided by anonymous communication systems, traffic analysis attacks have been studied [3, 7–14]. Existing traffic analysis attacks can be categorized into two groups: passive traffic analysis and active watermarking techniques. Passive traffic analysis technique will record the traffic passively and identify the correlation between sender’s outbound traffic and receiver’s inbound traffic based on statistical measures. This type of technique requires a relatively long period of traffic observation for a reasonable detection rate. The idea is to actively introduce special signals into the sender’s outbound traffic with the intention of recognizing the embedded signal at the receiver’s inbound traffic. Encryption does not work, since packet headers still reveal a great deal about users.

II. Anonymous Communication
Passive traffic analysis technique will record the traffic passively and identify the correlation between sender’s outbound traffic and receiver’s inbound traffic based on statistical measures. This type of technique requires a relatively long period of traffic observation for a reasonable detection rate. The idea is to actively introduce special signals into the sender’s outbound traffic with the intention of recognizing the embedded signal at the receiver’s inbound traffic. Encryption does not work, since packet headers still reveal a great deal about users.

We focus on the active watermarking technique, which has been active in the past few years. Proposed a flow-marking scheme based on the direct sequence spread spectrum technique by utilizing a pseudo-noise code. By interfering with the rate of a suspect sender’s traffic and marginally changing the traffic rate, the attacker can embed a secret spread-spectrum signal into the target traffic. The embedded signal is carried along with the target traffic from the sender to the receiver, so the investigator can recognize the corresponding communication relationship, tracing...
the messages despite the use of anonymous networks. However, in order to accurately confirm the anonymous communication relationship of users, the flow-marking scheme needs to embed a signal modulated by a relatively long length of PN code, and also the signal is embedded into the traffic flow rate variation. Houmansadr et al. proposed a no blind network flow watermarking scheme called RAINBOW for stepping stone detection.

Fig. 2: Architecture

A. Data Transmission
In Tor, a maintains a connection to other on demand. The uses a way of source routing and chooses several from the locally cached directory, downloaded from the directory caches. The number of the selected is referred as the path length. We use the default path length of three as an example. The iteratively establishes circuits across the Tor network and negotiates a symmetric key with each, one hop at a time, as well as handles the streams from client applications. The side of the circuit connects to the requested destinations and relays the data. We now illustrate the procedure that the establishes a circuit and downloads a file from the server.

B. Components of Tor
Onion routers are special proxies that relay the application data. In Tor, transport-layer security connections are used for the overlay link encryption between two onion routers. The application data is packed into equal-sized cells. They hold onion router information such as public keys for onion routers. Directory authorities hold authoritative information on onion routers, and directory caches download directory information of onion routers from authorities.

C. Cells at Onion Routers
To begin with, the onion router receives the data from the connection on the given port A. After the data is processed by protocols, the data will be delivered into the buffer of the connection. When there is pending data in the buffer, the read event of this connection will be called to read and process the data. The connection read event will pull the data from the buffer into the connection input buffer. Each connection input buffer is implemented as a linked list with small chunks. The data is fetched from the head of the list and added to the tail. After the data in the TLS buffer is pulled into the connection input buffer, the connection read event will process the cells from the connection input buffer one by one.

III. Applications
This application is mainly used in an internet applications where we would like implement in the anonymous which is the allows the attacker to confirm anonymous communication relationship among users very quickly. In this attack, by marginally varying the number of cells in the target traffic at the malicious exit onion router, the attacker can embed a secret signal into the variation of cell counter of the target traffic.

IV. Conclusion
Finally I would like to conclude that by interfering with the rate of a suspect sender’s traffic and marginally changing the traffic rate, the attacker can embed a secret spread-spectrum signal into the target traffic. The embedded signal is carried along with the target traffic from the sender to the receiver, so the investigator can recognize the corresponding communication relationship, tracing the messages despite the use of anonymous networks. However, in order to accurately confirm the anonymous communication relationship of users, the flow-marking scheme needs to embed a signal modulated by a relatively long length of PN code, and also the signal is embedded into the traffic flow rate variation.
References


